YCTP-P6-00
TMUP-HEL-0011

Aharonov-Bohm and Coulomb Scattering Near the Forward Direction[1]

Charles M. Sommerfield[2]
Department of Physics, Yale University,
New Haven, CT 06520-8120, USA

Hisakazu Minakata[3]
Department of Physics, Tokyo Metropolitan University,
Hachioji, Tokyo 192-0397, Japan
and
Research Center for Cosmic Neutrinos, Institute for Cosmic Ray Research,
University of Tokyo, Chiba 277-8582, Japan

June, 2000

Abstract

The exact wave functions that describe scattering of a charged particle by a confined magnetic field (Aharonov-Bohm effect) and by a Coulomb field are analyzed. It is well known that the usual procedure of finding asymptotic forms of these functions which admit a separation into a superposition of an incident plane wave and a circular or spherical scattered wave is problematic near the forward direction. It thus appears to be impossible to express the conservation of probability by means of an optical theorem of the usual kind. Both the total cross section and the forward scattering amplitude appear to be infinite. To address these difficulties we find a new representation for the asymptotic form of the Aharonov-Bohm wave function that is valid for all angles. Rather than try to define a cross section at forward angles, however, we work instead with the probability current and find that it is quite well behaved. The same is true for Coulomb scattering. We trace the usual difficulties to a nonuniformity of limits.

[1]Talk presented at Yale-TMU Symposium on Dynamics of Gauge Fields, Tokyo, Dec. 1999.

[2]charles.sommerfield@yale.edu

[3]minakata@phys.metro-u.ac.jp

1 Introduction

Not only is 1999 the fiftieth anniversary of Tokyo Metropolitan University, it also marks forty years since the publication of the paper of Aharonov and Bohm [1] in which it was shown that there are measurable effects that can be attributed directly to the electromagnetic vector potential and only non-locally to the magnetic field itself. Of course in non-Abelian gauge theories, which are the subject of this conference, the direct importance of the potential does not raise any eyebrows, but it did so with respect to electromagnetism at the time. The effect was soon confirmed experimentally by Chambers [2] and subsequently by many others [3]. In connection with anniversaries it may be interesting to note that in 1949 (again, fifty years ago) Ehrenberg and Siday published a paper [4] in which the ray optics of an electron moving near a magnetic field were analyzed. It is possible to view this work as a precursor to that of Aharonov and Bohm.

This paper by no means represents a comprehensive review. The literature of the subject is enormous. The reader can use the references we have provided to help trace its detailed history. We begin in Sec. 2 with a brief review of the Aharonov-Bohm effect. What is involved is the effect of the magnetic field on a charged-particle wave function if the particle is forbidden from entering the region where the field strength is not zero. We give a heuristic derivation of the influence of the field on the interference pattern of a double-slit experiment. Aharonov and Bohm showed that if the magnetic field is restricted to an infinite line and if the wave function of the particle is chosen to vanish on that line, then the problem is exactly soluble. Almost all work on the subject starts from their solution. The wave function is given as an infinite sum of partial waves. The bulk of our presentation consists of an analysis of this sum.

In Sec. 3 we apply the usual method of partial wave analysis to calculate the phase shifts. This also leads to a formal derivation of the optical theorem. However, since the phase shifts are constant for large partial waves the convergence of the procedure is suspect. In Sec. 4 we review other attempts to find the asymptotic form of the Aharonov-Bohm function and to reconcile the results with the optical theorem. Sec. 5 is devoted to a new presentation of the Aharonov-Bohm function. It has an asymptotic form which is valid at all angles. In Sec. 6 we present graphical results of the probability current distribution from which one usually calculates cross sections. The behavior near forward angles is quite finite and closely follows the double-slit interference pattern mentioned above.

In Sec. 7 we show that Coulomb scattering shares many of the same features of Aharonov-Bohm scattering. Again we provide graphical results of the probability current as derived from the exact Coulomb wave functions.

2 The Aharonov-Bohm effect and a simple model

Consider a charged particle in the presence of a magnetic field $\mathbf{B}(\mathbf{r})$ which is described by a vector potential $\mathbf{A}(\mathbf{r})$ such that $\mathbf{B} = \nabla \times \mathbf{A}$. If in the absence of the magnetic field the system is described by a stationary-state wave function $\psi_0(\mathbf{r})$, then in the presence of the field it will be described by the path-dependent multiple-valued wave function

$$\psi(\mathbf{r}) = \psi_0(\mathbf{r}) \exp\left[i\frac{q}{\hbar c}\int_{\mathbf{r}_0}^{\mathbf{r}} \mathbf{A}(\mathbf{r}') \cdot d\mathbf{r}'\right] = \psi_0(\mathbf{r}) e^{i\chi(\mathbf{r},P)}. \tag{1}$$

Here q is the charge of the particle and P symbolizes the path followed by the integral.

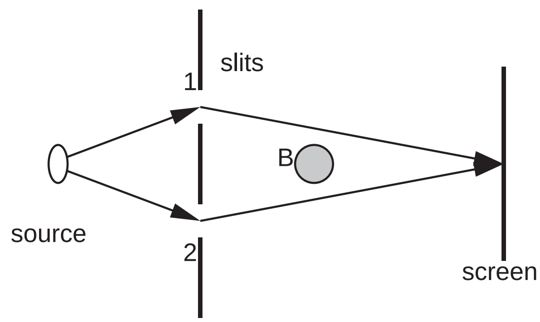

Figure 1: The Aharonov-Bohm double slit experiment.

If one arranges a double slit experiment, as in Fig. 1, so that the part of the beam that goes through slit 1 passes entirely above the magnetic field while the part that goes through slit 2 passes entirely below the field then using a ray-optics approximation one expects that at the particular position that would be the central point of the interference pattern in the absence of the field, the wave function will now be

$$\begin{aligned} \psi(\mathbf{r}) &= \tfrac{1}{2}\psi_0(\mathbf{r})[e^{i\chi(\mathbf{r},P_1)} + e^{i\chi(\mathbf{r},P_2)}] \\ &= \tfrac{1}{2}\psi_0(\mathbf{r})e^{i\chi(\mathbf{r},P_1)}\left[1 + \exp\left(i\frac{q\Phi}{\hbar c}\right)\right], \end{aligned} \tag{2}$$

where Φ is the flux of the magnetic field through the area bounded by the two paths. Thus the interference pattern will change with the magnetic field, even if the electron wave function is zero in the region to which the field is confined. This contradicts the classical notion that the vector potential is basically unphysical and that a charged particle can feel a force only at points where the magnetic field itself is nonzero.

The double slit geometry is not essential for the Aharonov-Bohm effect. Aharonov and Bohm treated an idealization of the situation without the slits that could be solved analytically. They restricted the magnetic field to be confined to a line of zero width oriented along the z direction. That is they took

$$\mathbf{B}(\mathbf{r}) = \hat{\mathbf{z}}\,\Phi\,\delta(x)\delta(y). \tag{3}$$

In this case, using plane polar coordinates r and θ in the x-y plane, one may take the vector potential to be

$$\mathbf{A}(r,\theta) = \frac{\Phi}{2\pi r}\hat{\boldsymbol{\theta}} = \frac{\Phi}{2\pi}\nabla\theta. \tag{4}$$

The many-valued wave function in the presence of the field would be simply

$$\psi(r,\theta) = \psi_0(r,\theta)\exp\left[i\frac{q}{\hbar c}\frac{\Phi}{2\pi}\theta\right] = \psi_0(\theta)e^{i\alpha\theta}. \tag{5}$$

We have defined α to measure the strength of the flux. It is convenient for later developments to take the greatest integer in α to be N and to write

$$\alpha = \frac{q}{\hbar c}\frac{\Phi}{2\pi} = [\alpha] + \tfrac{1}{2} + \gamma = N + \tfrac{1}{2} + \gamma, \tag{6}$$

with $-\frac{1}{2} < \gamma < \frac{1}{2}$.

The Aharonov-Bohm scattering problem is to find a single-valued solution of Schrödinger's equation that corresponds as closely as possible to an incident plane wave moving to the left along the x axis from the right of the magnetic field in Fig. 1 (with the source, screen and slits eliminated) subject to the boundary condition that the wave function vanishes wherever the field is nonzero. For the field of Eq. (3) we must take the wave function to vanish at the origin. Ignoring the z coordinate, the Schrödinger equation is

$$\left[\left(\nabla - i\frac{q}{\hbar c}\mathbf{A}(\mathbf{r})\right)^2 + k^2\right]\psi(r,\theta) = 0, \tag{7}$$

which simplifies to

$$\left[\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\left(\frac{\partial}{\partial\theta} - i\alpha\right)^2 + k^2\right]\psi(r,\theta) = 0. \tag{8}$$

The wave function in Eq. (5), which in this case is $\psi(r,\theta) = e^{-ikr\cos\theta}e^{i\alpha\theta}$, is clearly not acceptable despite the fact it solves Eq. (8), since it does not vanish at the origin and is not single-valued. A single-valued solution can be found by a partial wave expansion and becomes a superposition over all l of solutions $J_{l-\alpha}(kr)e^{il\theta}$ that are eigenfunctions of the z component of angular momentum. The wave function of Aharonov and Bohm is of this form and does reduce to the simple plane wave when $\alpha = 0$. It is

$$\psi_{\mathrm{AB}}(r,\theta) = \sum_{l=-\infty}^{\infty} e^{-i\pi|l-\alpha|/2} J_{|l-\alpha|}(kr)e^{il\theta}. \tag{9}$$

Using a concept he calls "whirling waves," Berry [5] has given a beautiful derivation of this result that reconciles the requirement of single-valuedness with the notion that the effect of the magnetic field can be given by a simple multiplicative phase factor.

3 Two-dimensional partial-wave expansion

The asymptotic form for large r of a wave function such as that given in Eq. (9) is expected to have the form

$$\psi_{\rm AB}(r,\theta) \to e^{-ikr\cos\theta} + e^{i\frac{\pi}{4}} f(\theta)\frac{e^{ikr}}{\sqrt{r}}, \tag{10}$$

where the scattering amplitude $f(\theta)$ would have the partial-wave expansion

$$f(\theta) = \sqrt{\frac{2}{\pi k}} \sum_{l=-\infty}^{\infty} (-1)^l e^{i\delta_l} \sin\delta_l \, e^{il\theta}. \tag{11}$$

An examination of Eq. (9) shows the phase shifts to be $\delta_l = \frac{\pi}{2}(l - |l-\alpha|)$ so that

$$f(\theta) = \sqrt{\frac{2}{\pi k}} \sin\frac{\pi}{2}\alpha \left[e^{i\frac{\pi}{2}\alpha} \sum_{l=N+1}^{\infty} (-1)^l e^{il\theta} e^{-2\epsilon l} - e^{-i\frac{\pi}{2}\alpha} \sum_{l=-\infty}^{N} (-1)^l e^{il\theta} e^{2\epsilon l} \right], \tag{12}$$

where convergence factors involving the small positive quantity ϵ have been inserted to insure convergence. The sums are easily done and the result is

$$f(\theta) = \frac{1}{\sqrt{2\pi k}} (-1)^N e^{i(N+\frac{1}{2})\theta} \left[\frac{-\sin\pi\alpha\cos\frac{\theta}{2} + (1-\cos\pi\alpha)\epsilon\sin\frac{\theta}{2}}{\cos^2\frac{\theta}{2} + \epsilon^2} \right]. \tag{13}$$

Letting $\epsilon \to 0$ we find

$$f(\theta) = -\frac{\sin\pi\alpha}{\sqrt{2\pi k}} (-1)^N e^{i(N+\frac{1}{2})\theta} \frac{1}{\cos\frac{\theta}{2}} + i\pi \frac{1-\cos\pi\alpha}{\sqrt{2\pi k}} \delta(\cos\tfrac{\theta}{2}), \tag{14}$$

which is a result obtained by Ruijsenaars [6]. Criticism of this formal method of summing the partial-wave expansion has been given by Hagen [7]. Arai and Minakata [8] have analyzed this expression for $f(\theta)$ in connection with the optical theorem in great detail without a completely satisfactory conclusion. On the other hand Sakoda and Omote have shown that the S-matrix is indeed unitary [9].

4 Integral representations of the Aharonov-Bohm wave function

Another method that makes sense of the sum in Eq. (9) is to use a modified version of Sommerfeld's representation of the Bessel function

$$J_\nu(z) = \frac{1}{2\pi} e^{i\nu\frac{\pi}{2}} \int_{\mathcal{C}} dt \, e^{-iz\cos t} e^{i\nu t}, \tag{15}$$

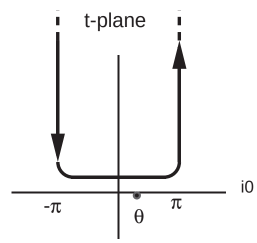

Figure 2: The contour $\mathcal{C}$ of Eq. (15).

where $\mathcal{C}$ goes from $-\pi+\eta+i\infty$ to $\pi+\eta+i\infty$ with η a positive infinitesimal as illustrated in Fig. 2.

When this representation is used in Eq. (9) one can perform the sum over l under the integral sign to obtain

$$\psi_{\rm AB}(r,\theta) = \frac{i}{4\pi} e^{(N+\frac{1}{2})\theta} \int_{\mathcal{C}} dt\; e^{-ikr\cos t} \left[\frac{e^{i\gamma t}}{\sin\frac{1}{2}(t-\theta)} + \frac{e^{-i\gamma t}}{\sin\frac{1}{2}(t+\theta)} \right]. \tag{16}$$

One way to proceed is to note that the substitution $t \to -t$ in the second term of the integral in Eq. (16) leads to an expression identical with the first term but integrated now along a contour that is $\mathcal{C}$ reflected in the origin. The contours may be distorted, as shown by comparison of the left- and right-hand illustrations of Fig. 3, into two straight-line paths and a clockwise contour around the pole at $t=\theta$.

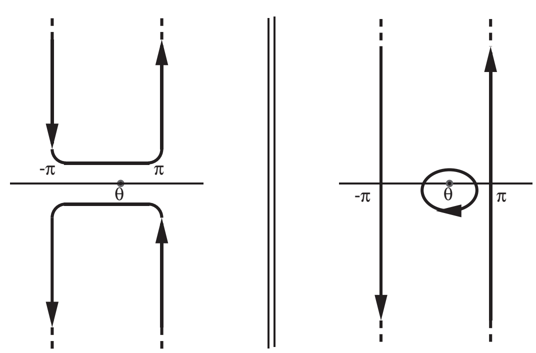

Figure 3: Integration contours for Eq. (16).

The contour around the pole gives rise to a modified incident wave of the form $e^{-ikr\cos\theta}e^{i\alpha\theta}$ which is not single-valued. The other terms may be analyzed by steepest-descent methods and give rise to a scattered wave that is valid except for θ near $\pm\pi$. The result is

$$\psi_{\rm AB}(r,\theta) \to e^{-ikr\cos\theta}e^{i\alpha\theta} - e^{i\frac{\pi}{4}} \frac{e^{ikr}}{\sqrt{r}} \frac{1}{\sqrt{2\pi k}} (-1)^N \sin\alpha\pi \frac{e^{i(N+\frac{1}{2})\theta}}{\cos\frac{1}{2}\theta}. \tag{17}$$

Although the scattered wave is not single-valued, the entire wave function is. It remains unclear from this approach, however, how to handle scattering near the forward direction.

5 A new integral representation

We go back to the original contour $\mathcal{C}$ and write

$$\psi_{\rm AB}(r,\theta) = \frac{i}{4\pi} e^{i(N+\frac{1}{2})\theta} e^{-ikr\cos\theta} A(kr,\theta), \tag{18}$$

which, by comparison with Eq. (16) defines $A(kr,\theta)$. By differentiating $A(kr,\theta)$ with respect to kr we can eliminate the singularities. We find, after some manipulation of the resulting Bessel functions

$$\begin{aligned}\frac{\partial}{\partial(kr)} A(kr,\theta) &= -2\pi i \cos\gamma\pi\, e^{ikr\cos\theta} \\ &\times \left[e^{i\frac{1}{2}\theta} e^{i(\frac{1}{2}+\gamma)\frac{\pi}{2}} H^{(1)}_{\frac{1}{2}+\gamma}(kr) + e^{-i\frac{1}{2}\theta} e^{i(\frac{1}{2}-\gamma)\frac{\pi}{2}} H^{(1)}_{\frac{1}{2}-\gamma}(kr) \right],\end{aligned} \tag{19}$$

in which Hankel functions appear. Since the wave function must vanish for $r = 0$ we integrate this result from 0 to kr and obtain the following exact expression for the Aharonov-Bohm wave function.

$$\begin{aligned}\psi_{\rm AB}(r,\theta) &= \tfrac{1}{2} e^{-i\frac{\pi}{4}} e^{i(N+\frac{1}{2})\theta} \cos\gamma\pi\, e^{-ikr\cos\theta} \\ &\times \int_0^{kr} dz\, e^{iz\cos\theta} \left[e^{i\frac{1}{2}\theta} e^{i\gamma\frac{\pi}{2}} H^{(1)}_{\frac{1}{2}+\gamma}(z) + e^{-i\frac{1}{2}\theta} e^{-i\gamma\frac{\pi}{2}} H^{(1)}_{\frac{1}{2}-\gamma}(z) \right].\end{aligned} \tag{20}$$

It should be noted that this expression is valid only for $-\pi \le \theta \le \pi$. For other values of θ one should replace θ by $[(\theta+\pi)\mathrm{mod}(2\pi)] - \pi$.

The asymptotic form of $\psi_{\rm AB}(r,\theta)$ for large r and any value of θ may be found by writing the integrals from 0 to kr in Eq. (20) as the difference of integrals from 0 to ∞ and from kr to ∞. For large kr we may use the large-argument asymptotic form of the Hankel functions to approximate the second integral (no matter what the value of θ) while the first integral can be done exactly [10]. The result is

$$\psi_{\rm AB}(r,\theta) \to e^{i(N+\frac{1}{2})\theta} e^{-ikr\cos\theta} \left\{ e^{i\gamma\theta} - \cos\gamma\pi \left[1 - \mathrm{erf}\left(e^{-i\frac{\pi}{4}} \sqrt{2kr} \cos\tfrac{1}{2}\theta \right) \right] \right\}. \tag{21}$$

Stelitano [11] has obtained a similar expression by differentiating the partial-wave sum directly with respect to kr. But he treats the $l = 0$ term separately which means that his asymptotic expression is not as compact as Eq. (21) nor is it strictly correct for backward scattering angles. Alvarez [12] has also given an asymptotic form for large r as an infinite series that is applicable for any θ.

For $\sqrt{kr}\cos\frac{1}{2}\theta \gg 1$, (non-forward angles), Eq. (21) reduces to Eq. (17), in agreement with the other integration method, while for $\theta = \pi$ and large kr we get $\psi(r,\pi) \to \cos\alpha\pi\, e^{ikr}$, which is perfectly finite and which exhibits the Aharonov-Bohm interference.

In view of these results it is to be stressed once again, as has been done by many previous workers on this problem, that near the forward direction the total wave function does not admit a separation into incident and scattered parts. Hence the usual statement of the optical theorem as an expression of unitarity just doesn't make sense.

6 Properties of the probability current

From the time-independent point of view the usual optical theorem reflects the conservation of the probability current, the integrated form of which is $\int_{-\pi}^{\pi} d\theta\, \hat{\mathbf{r}} \cdot \mathbf{j}(\mathbf{r}) = 0$, for any fixed r. The probability current density is given by

$$\mathbf{j}(\mathbf{r}) = \frac{\hbar}{m} \mathrm{Im}[\psi^*(\mathbf{r}) \nabla \psi(\mathbf{r})]. \tag{22}$$

It usually has three terms—coming from the incident wave, the scattered wave and the interference between them. In a wave-packet treatment the latter has significance only near the forward direction. Because it is proportional to $1/r$ the scattered term will be quite small compared to the incident term. And even at forward angles it will be overwhelmed by the incident term.

The large kr approximation given by Eq. (21) greatly simplifies the numerical calculation of $\hat{\mathbf{r}} \cdot \mathbf{j}(\mathbf{r})$. The results are plotted in Fig. 4 for $\gamma = .25$ and $kr = 100$, while in Fig. 5 we show the naively defined cross section for the same parameters. What we see in the Aharonov-Bohm case, where we cannot make the separation into three terms, is that effectively the forward scattering becomes significant enough to completely modify the behavior in the forward direction, while the probability current remains perfectly well-behaved and continuous there.

It is clear from this analysis that the limits of $\theta \to \pi$ and $r \to \infty$ do not commute. It is our contention that the former limit is what makes sense physically for any finitely contained

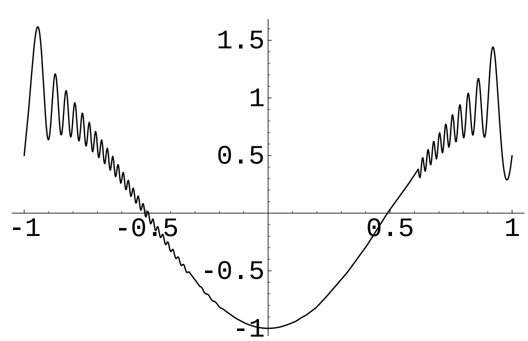

Figure 4: A plot of $\hat{\mathbf{r}} \cdot \mathbf{j}(\mathbf{r})$ against θ/π for $kr = 100$ and $\gamma = .25$, normalized to -1 at $\theta = 0$.

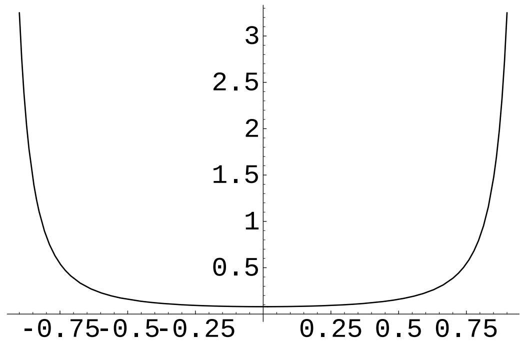

Figure 5: A plot of the usual differential cross section in units of $1/k$ against θ/π for $kr = 100$ and $\gamma = .25$.

experiment.

7 Coulomb scattering

A similar nonuniformity of limits appears in the consideration of Coulomb scattering. The exact wave function, written in terms of a confluent hypergeometric function is

$$\psi_{\text{Coul}}(\mathbf{r}) = e^{ikr\cos\theta}\ {}_1F_1(-i\eta; 1; ikr[1-\cos\theta]), \tag{23}$$

with Coulomb parameter $\eta = mZe^2/(\hbar k)$. For large $kr(1-\cos\theta)$ this has the asymptotic form

$$\psi_{\text{Coul}}(\mathbf{r}) \sim e^{i\eta\log kr(1-\cos\theta)}\left[e^{ikr\cos\theta} - \frac{e^{ikr}}{kr}\frac{e^{-2i\eta\log kr}}{(1-\cos\theta)^{1+2i\eta}}\frac{\Gamma(1+i\eta)}{\Gamma(1-i\eta)}\eta\right], \tag{24}$$

which looks like a plane wave incident along the z axis plus a spherical wave, but with Coulomb distortion. The "r-dependent" scattering amplitude extracted from this is

$$f(\theta) = -\frac{\eta}{k}\frac{\Gamma(1+i\eta)}{\Gamma(1-i\eta)}\frac{e^{-2i\eta\log kr}}{(1-\cos\theta)^{1+2i\eta}} \tag{25}$$

and the differential cross section is

$$\frac{d\sigma}{d\Omega} = \frac{\eta^2}{k^2(1-\cos\theta)^2}. \tag{26}$$

When extrapolated to forward angles these are both infinite.

On the other hand, the exact wave function and the probability current are quite finite at $\theta = 0$. In fact the probability current becomes very small compared to its maximum value at near-forward angles. This is illustrated in Fig. 6 with $\eta = 1.5$.

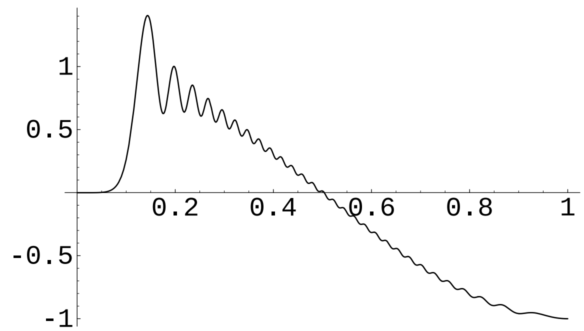

Figure 6: Radial component of the probability current for Coulomb scattering, plotted against θ/π and normalized to -1 at backward angles. (The forward angle is now at $\theta = 0$). The parameters are $kr = 100$ and $\eta = 1.5$.

Acknowledgments

This work was supported in part by the U.S. Department of Energy under contract No. DE-FG-02-92ER-10701 and by Grant-in-Aid for Scientific Research No. 09045036 under the International Scientific Research Program, Inter-University Cooperative Research, Japanese Ministry of Education, Science, Sports and Culture.

References

[1] Y. Aharonov and D. Bohm, Phys. Rev. 115 (1959) 485.

[2] R. G. Chambers, Phys. Rev. Lett. 5 (1960) 3.

[3] A. Tonomura et al., Phys. Rev. Lett. 56 (1986) 792.

[4] W. Ehrenberg and R. E. Siday, Proc. Phys. Soc. (London) B62 (1949) 8.

[5] M. Berry, Eur. J. Phys. 1 (1980) 240.

[6] S. N. M. Ruijsenaars, Ann. Phys. (N.Y.) 146 (1983) 1.

[7] C. Hagen, Phys. Rev. D 41 (1990) 2015.

[8] M. Arai and H. Minakata, Intl. J. Mod. Phys. A13 (1998) 831, `hep-th/9610219`.

[9] S. Sakoda and M. Omote, J. Math. Phys. 38 (1997) 716, `hep-th/9603185`.

[10] "Handbook of Mathematical Functions," M. Abramowitz and I. Stegun eds., p. 487.

[11] D. Stelitano, Phys. Rev. D 51 (1995) 5876, hep-th/9411175.

[12] M. Alvarez, Phys. Rev. A 54 (1996), 1128, hep-th/9510085.